

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,994

DATE: 06/07/2001

TIME: 09:22:21

Input Set : A:\DES.txt

Output Set: C:\CRF3\06072001\I830994.raw

3 <110> APPLICANT: YLIHONKO, Kristiina  
 4 RAETY, Kaj  
 5 HAKALA, Juha  
 7 <120> TITLE OF INVENTION: THE GENE CLUSTER INVOLVED IN ACLACINOMYCIN BIOSYNTHESIS, AND  
 ITS USE FOR  
 8 GENETIC ENGINEERING  
 10 <130> FILE REFERENCE: 1574/49884  
 C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/830,994  
 C--> 12 <141> CURRENT FILING DATE: 2000-09-25  
 12 <150> PRIOR APPLICATION NUMBER: PCT/FI00/00819  
 13 <151> PRIOR FILING DATE: 2000-09-25  
 15 <160> NUMBER OF SEQ ID NOS: 16  
 17 <170> SOFTWARE: PatentIn version 3.0  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 662  
 21 <212> TYPE: PRT  
 22 <213> ORGANISM: Streptomyces galilaeus  
 24 <400> SEQUENCE: 1  
 26 Met Thr Glu Asp Arg Val Thr Thr Leu Gly Gly Glu Gln Ile Ala Leu  
 27 1 5 10 15  
 29 Leu Ala Pro Leu Leu Asp Gly Ser Arg Asp Leu Pro Gly Ile Val Ala  
 30 20 25 30  
 32 Asp Ala Ala Pro Arg Leu Pro Ala Gly Leu Ala Glu Arg Leu Val Thr  
 33 35 40 45  
 35 Arg Leu Leu Asp Ala Gly Leu Leu Cys Ala Tyr Pro Gln Asp Gly Ala  
 36 50 55 60  
 38 Asp Arg Pro Glu Arg Ala Tyr Arg Ser Leu Thr Gly Leu Gln Ala Arg  
 39 65 70 75 80  
 41 Ser Ala Asp Ala Arg Asp Ala Val Leu Ala Ala Val Asp Leu Thr Gly  
 42 85 90 95  
 44 Asp Ala Glu Ser Pro Leu Pro Glu Ala Val Ser Ala Ala Gly Leu Arg  
 45 100 105 110  
 47 Ala Ala Ala Pro Gly Glu His Ala Ala Leu Thr Leu Val Leu Cys His  
 48 115 120 125  
 50 Asp Tyr Leu Asp Pro Arg Leu Ser Ala Leu Asp Ala Glu His Arg Ala  
 51 130 135 140  
 53 Thr Gly Arg Gly Trp Leu Pro Val Arg Ala Asn Gly Thr His Leu Trp  
 54 145 150 155 160  
 56 Ile Gly Pro Phe Phe Ser Ala Gly Asp Gly Pro Cys Trp Ser Cys Leu  
 57 165 170 175  
 59 Ala Asp Arg Leu Arg Leu Arg Arg Arg Gly Glu Ala Tyr Val Gln His  
 60 180 185 190  
 62 Arg Leu Gly His Ser Gly Pro Ala Val His Arg Arg Ala Tyr Leu Pro  
 63 195 200 205  
 65 Ala Gly Arg Ala Ala Ala Leu Gln Leu Ala Leu Leu Glu Ala Gly Lys  
 66 210 215 220  
 68 Trp Leu Ser Gly His Arg Asp Thr Val Gln Asp Ser Leu Trp Arg Leu  
 69 225 230 235 240

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,994

DATE: 06/07/2001

TIME: 09:22:21

Input Set : A:\DES.txt

Output Set: C:\CRF3\06072001\I830994.raw

```

71 Asp Thr Arg Thr Leu Glu Ser Ser Arg His Pro Val Arg Arg Arg Pro
72                245                250                255
74 Gln Cys Ser Arg Cys Gly Asp Pro Leu Leu Val Arg Asp Arg Val Ser
75                260                265                270
77 Ala Pro Val Val Leu Ser Ser Arg Pro Val Arg Asp Glu Ser Gly Gly
78                275                280                285
80 Gly His Arg Thr Phe Gly Pro Gln Glu Met Leu Asp Arg Tyr Gly His
81                290                295                300
83 Leu Val Asp Pro Val Thr Gly Val Val Gly Glu Ile Arg Arg Asp Pro
84 305                310                315                320
86 Arg Gly Pro Glu Phe Leu Asn Cys Phe Thr Arg Ser Arg Cys Arg Leu
87                325                330                335
89 Gly Pro Arg Ala Ala Pro Pro Ala Leu His Ser Pro Leu Arg Ser Pro
90                340                345                350
92 Gly Ser Gly Lys Gly Val Thr Glu Leu His Ala Arg Val Ser Ala Leu
93                355                360                365
95 Ala Glu Ala Leu Glu Arg Cys Ser Gly Tyr Phe Gln Gly Asp Glu Pro
96                370                375                380
98 Arg Arg Arg Gly Ser Tyr Arg Glu Leu Ala Gly Leu Ala Val His Pro
99 385                390                395                400
101 Asp Ser Val Gln Leu Phe Asp Arg Arg Gln Phe Glu Asp Arg Arg Ala
102                405                410                415
104 Trp Asn Arg Ala His Gly Pro Phe His Gln Val Thr Glu Pro Phe Asp
105                420                425                430
107 Glu Asp Ala Pro Ile Asp Trp Thr Pro Val Trp Ser Leu Thr Glu Arg
108                435                440                445
110 Arg Gln Arg Leu Ala Pro Thr Ser Leu Leu Tyr Tyr Asn Ala Pro Asp
111                450                455                460
113 Ala Asp Thr Gly Phe Cys Arg Ala Thr Ser Asn Gly Ala Ala Ala Gly
114 465                470                475                480
116 Thr Ser Leu Glu Asp Ala Val Val His Gly Cys Leu Glu Leu Val Glu
117                485                490                495
119 Arg Asp Ala Ile Ala Leu Trp Trp Tyr Asn Arg Thr Arg Gln Pro Gly
120                500                505                510
122 Val Thr Leu Asp Ala Arg Asp Pro Trp Ile Thr Arg Leu Arg Ala Val
123                515                520                525
125 Leu Arg Asp Leu Gly Arg Thr Val Trp Ala Leu Asp Leu Thr Ser Asp
126                530                535                540
128 Leu Gly Ile Pro Val Val Ala Ala Val Ser Val Arg Thr Gly Gly Thr
129 545                550                555                560
131 Ala Glu Asp Ile Val Leu Gly Phe Gly Ala His Phe Asp Pro Arg Ile
132                565                570                575
134 Ala Leu Arg Arg Ala Leu Thr Glu Leu Ser Gln Met Leu Pro Pro Leu
135                580                585                590
137 Ala Gln Glu Thr Ala Gly Asp Ala Ser Ala Tyr Thr Gly Thr Asp Pro
138                595                600                605
140 Glu Ala Met Arg Trp Phe Arg His Ala Thr Thr Ala Asn Gln Pro Tyr
141                610                615                620
143 Leu Leu Pro Ala Ala Arg Arg Ser Ala Arg Pro Pro Ala Ser Leu Arg

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,994

DATE: 06/07/2001

TIME: 09:22:21

Input Set : A:\DES.txt

Output Set: C:\CRF3\06072001\I830994.raw

```

144 625                      630                      635                      640
146 Pro Pro Arg Asp Ala Ala Ala Gln Ala Gly Ala Leu Val Ala Leu Leu
147                      645                      650                      655
149 Arg Arg His Gly Leu Glu
150                      660
152 <210> SEQ ID NO: 2
153 <211> LENGTH: 272
154 <212> TYPE: PRT
155 <213> ORGANISM: Streptomyces galilaeus
157 <400> SEQUENCE: 2
159 Val Asp Ile Trp Leu Leu Gly Pro Leu Thr Ala Glu Val Arg Gly Arg
160 1                      5                      10                      15
162 Ser Ile Val Pro Thr Ala Ala Lys Pro Arg Gln Ile Leu Ala Leu Leu
163                      20                      25                      30
165 Ala Ile His Ala Asn Arg Val Leu Pro Val Gly Thr Leu Met Glu Glu
166                      35                      40                      45
168 Ile Trp Gly Thr Glu Pro Pro Gln Ser Ala Leu Ala Thr Leu His Thr
169                      50                      55                      60
171 Tyr Ile Leu Gln Leu Arg Arg Arg Leu Thr Ala Ala Tyr Gly Asp Glu
172 65                      70                      75                      80
174 Gly Gly Val Ser Ala Lys Asp Val Leu Val Thr Gln Tyr Gly Gly Tyr
175                      85                      90                      95
177 Cys Trp Gln Ala Pro Thr Asp Ser Val Asp Val Pro Arg Tyr Glu Arg
178                      100                     105                     110
180 Leu Val Thr Ala Gly Arg Ile Ala Thr Ala Glu Asp Arg Gln Glu Glu
181                      115                     120                     125
183 Ala Ser Ala His Phe Arg Glu Ala Leu Ala Leu Trp Arg Gly Ser Ala
184                      130                     135                     140
186 Leu Val Asp Val Arg Ile Gly Pro Val Leu Ser Ile Glu Val Ala Arg
187 145                     150                     155                     160
189 Leu Glu Glu Ser Arg Leu Gly Val Leu Glu Arg Cys Leu Glu Ala Asp
190                      165                     170                     175
192 Leu Arg Leu Gly Arg His Ala Glu Leu Leu Ala Glu Leu Thr Glu Leu
193                      180                     185                     190
195 Thr Gly Arg His Pro Leu His Glu Gly Leu His Ala Gln Cys Met Thr
196                      195                     200                     205
198 Ala Leu Tyr Arg Ala Gly Arg Ser Trp Gln Ala Leu Asp Val Tyr Gln
199                      210                     215                     220
201 Arg Leu Arg Arg Arg Leu Ala Glu Glu Leu Gly Leu Ser Pro Ser Pro
202 225                     230                     235                     240
204 Arg Leu Gln Arg Leu Gln Gln Ala Val Leu Ser Ala Glu Pro Trp Leu
205                      245                     250                     255
207 Asp Ala Pro Arg Tyr Gly Gly Asp Pro Val Phe Asp Arg Met Ile Ser
208                      260                     265                     270
210 <210> SEQ ID NO: 3
211 <211> LENGTH: 434
212 <212> TYPE: PRT
213 <213> ORGANISM: Streptomyces galilaeus
215 <400> SEQUENCE: 3

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,994

DATE: 06/07/2001

TIME: 09:22:21

Input Set : A:\DES.txt

Output Set: C:\CRF3\06072001\I830994.raw

```

217 Met Thr Ser Asp Thr Lys Ala Leu Val Leu Glu Gln Val Arg Glu Tyr
218 1 5 10 15
220 His Arg Gln Gln Gln Pro Gly Asn Phe Gln Pro Gly Val Thr Pro Ile
221 20 25 30
223 Leu Ser Ser Gly Ala Val Leu Asp Glu Glu Asp Arg Val Ala Leu Val
224 35 40 45
226 Glu Ala Ala Leu Asp Leu Arg Ile Ala Ala Gly Ala His Ser Arg Arg
227 50 55 60
229 Phe Glu Ser Lys Phe Ala Arg His Ile Gly Val Arg Lys Ala His Leu
230 65 70 75 80
232 Val Asn Ser Gly Ser Ser Ala Asn Leu Leu Ala Leu Ser Ala Leu Thr
233 85 90 95
235 Ser Pro Arg Leu Gly Glu Gln Arg Leu Arg Pro Gly Asp Glu Val Ile
236 100 105 110
238 Thr Val Ala Gly Gly Phe Pro Thr Thr Val Asn Pro Ile Leu Gln Asn
239 115 120 125
241 Gly Leu Thr Pro Val Phe Val Asp Leu Glu Leu Gly Thr Tyr Asn Thr
242 130 135 140
244 Thr Val Glu His Val Arg Ala Ala Ile Ser Asp Arg Thr Arg Ala Ile
245 145 150 155 160
247 Met Ile Ala His Thr Leu Gly Asn Pro Tyr Gln Val Ala Glu Ile Gln
248 165 170 175
250 Gln Leu Ala Thr Glu His Glu Leu Phe Leu Ile Glu Asp Asn Cys Asp
251 180 185 190
253 Ala Val Gly Ser Thr Tyr Gln Gly Arg Met Thr Gly Thr Phe Gly Asp
254 195 200 205
256 Leu Ala Thr Val Ser Phe Tyr Pro Ala His His Ile Thr Thr Gly Glu
257 210 215 220
259 Gly Gly Cys Val Leu Thr Arg Asn Leu Glu Leu Ala Arg Ile Val Glu
260 225 230 235 240
262 Ser Phe Arg Asp Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Glu Asp
263 245 250 255
265 Asn Thr Cys Leu Lys Arg Phe Asp Tyr Gln Leu Gly Asn Leu Pro Lys
266 260 265 270
268 Gly Tyr Asp His Lys Tyr Ile Phe Ser His Ile Gly Tyr Asn Leu Lys
269 275 280 285
271 Ala Thr Asp Leu Gln Gly Ala Leu Ala Leu Ser Gln Leu Asn Lys Leu
272 290 295 300
274 Pro Glu Phe Gly Ala Ala Arg Arg Arg Asn Trp Gln Arg Leu Arg Asp
275 305 310 315 320
277 Gly Leu Ala Asp Val Pro Gly Leu Leu Leu Pro Val Ala Thr Pro Gly
278 325 330 335
280 Ser Asp Pro Ser Trp Phe Gly Phe Val Ile Thr Val Leu Pro Asp Ala
281 340 345 350
283 Thr Tyr Thr Arg Arg Asp Leu Val Ala Phe Leu Glu Glu Arg Arg Ile
284 355 360 365
286 Gly Thr Arg Arg Leu Phe Gly Gly Asn Leu Thr Arg His Pro Ala Tyr
287 370 375 380
289 Leu Gly Thr Pro His Arg Val Ala Gly Asp Leu Arg Asn Ser Asp Ile

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/830,994

DATE: 06/07/2001

TIME: 09:22:21

Input Set : A:\DES.txt

Output Set: C:\CRF3\06072001\I830994.raw

```

290 385          390          395          400
292 Ile Thr Glu Gln Ser Phe Trp Ile Gly Val Tyr Pro Gly Ile Thr Glu
293          405          410          415
295 Glu Met Thr Asp Tyr Met Arg Glu Ser Ile Val Glu Phe Val Thr Lys
296          420          425          430
298 Asn Gly
301 <210> SEQ ID NO: 4
302 <211> LENGTH: 329
303 <212> TYPE: PRT
304 <213> ORGANISM: Streptomyces galilaeus
306 <400> SEQUENCE: 4
308 Met Pro Lys Asp Thr Pro Arg Pro Val Leu Arg Ile Gly Val Leu Gly
309 1          5          10          15
311 Cys Ala Asp Ile Ala Val Arg Arg Ile Leu Pro Ala Ile Val Glu His
312          20          25          30
314 Pro Ser Val Arg Leu Val Ala Leu Ala Ser Arg Asp Gly Ala Arg Ala
315          35          40          45
317 Glu Arg Leu Ala Ala Arg Phe Gly Cys Ala Ala Val Thr Gly Tyr Lys
318          50          55          60
320 Ala Leu Leu Asp Arg Glu Asp Ile Asn Ala Val Tyr Val Pro Leu Pro
321 65          70          75          80
323 Pro Gly Met His His Glu Trp Val Thr Glu Ala Leu Thr Ala Gly Lys
324          85          90          95
326 His Val Leu Val Glu Lys Pro Leu Ser Thr Thr Tyr Ala Gln Ser Val
327          100          105          110
329 Asp Leu Val Ala Met Ala Gly Arg Leu Gly Leu Ala Leu Thr Glu Asn
330          115          120          125
332 Phe Met Phe Leu His His Ser Gln His Glu Ala Val Arg Ala Met Thr
333          130          135          140
335 Gly Glu Ile Gly Glu Leu Arg Val Phe Thr Ser Ser Phe Gly Val Pro
336 145          150          155          160
338 Pro Pro His Pro Ser Ser Phe Arg His Asp Ala Arg Leu Gly Gly Gly
339          165          170          175
341 Ala Leu Leu Asp Val Gly Val Tyr Pro Leu Arg Ala Ala Gln Leu His
342          180          185          190
344 Leu Ala Gly Glu Leu Asp Val Leu Gly Ala Cys Leu Arg Val Asp Glu
345          195          200          205
347 Ala Thr Gly Val Asp Val Ala Gly Ser Ala Leu Leu Ser Thr Ala Thr
348          210          215          220
350 Gly Val Thr Ala Gln Leu Asp Phe Gly Phe Gln His Ala Tyr Arg Ser
351 225          230          235          240
353 Val Tyr Ala Leu Trp Gly Ser Arg Gly Arg Leu Ser Val Pro Arg Ala
354          245          250          255
356 Phe Thr Pro Pro Arg Glu His Arg Pro Val Val Arg Ile Glu Gln Gln
357          260          265          270
359 Asp Arg Leu Thr Glu Val Thr Leu Pro Ala Asp His Gln Val Gly Asn
360          275          280          285
362 Ala Leu Asp Ala Phe Ala Ser Ala Val His Ser Glu Thr Val Arg Ala
363          290          295          300

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

**VERIFICATION SUMMARY**PATENT APPLICATION: **US/09/830,994**

DATE: 06/07/2001

TIME: 09:22:22

Input Set : **A:\DES.txt**Output Set: **C:\CRF3\06072001\I830994.raw**

L:12 M:270 C: Current Application Number differs, Replaced Current Application No  
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:1168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14  
L:1472 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15  
L:1485 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16